

12th Workshop on Asian Language Resources 2016 (ALR 12)

Osaka, Japan
12 December 2016

ISBN: 978-1-5108-3334-0

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© (2016) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2016)

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: 845-758-2633
Email: curran@proceedings.com
Web: www.proceedings.com

Table of Contents

<i>An extension of ISO-Space for annotating object direction</i>	
Daiki Gotou, Hitoshi Nishikawa and Takenobu Tokunaga	1
<i>Annotation and Analysis of Discourse Relations, Temporal Relations and Multi-Layered Situational Relations in Japanese Texts</i>	
Kimi Kaneko, Saku Sugawara, Koji Mineshima and Daisuke Bekki	10
<i>Developing Universal Dependencies for Mandarin Chinese</i>	
Herman Leung, Rafaël Poiret, Tak-sum Wong, Xinying Chen, Kim Gerdes and John Lee	20
<i>Developing Corpus of Lecture Utterances Aligned to Slide Components</i>	
Ryo Minamiguchi and Masatoshi Tsuchiya	30
<i>VSoLSCSum: Building a Vietnamese Sentence-Comment Dataset for Social Context Summarization</i>	
Minh-Tien Nguyen, Dac Viet Lai, Phong-Khac Do, Duc-Vu Tran and Minh-Le Nguyen	38
<i>BCCWJ-DepPara: A Syntactic Annotation Treebank on the ‘Balanced Corpus of Contemporary Written Japanese’</i>	
Masayuki Asahara and Yuji Matsumoto	49
<i>SCTB: A Chinese Treebank in Scientific Domain</i>	
Chenhui Chu, Toshiaki Nakazawa, Daisuke Kawahara and Sadao Kurohashi	59
<i>Big Community Data before World Wide Web Era</i>	
Tomoya Iwakura, Tetsuro Takahashi, Akihiro Ohtani and Kunio Matsui	68
<i>An Overview of BPPT’s Indonesian Language Resources</i>	
Gunarso Gunarso and Hammam Riza	73
<i>Creating Japanese Political Corpus from Local Assembly Minutes of 47 prefectures</i>	
Yasutomo Kimura, Keiichi Takamaru, Takuma Tanaka, Akio Kobayashi, Hiroki Sakaji, Yuzu Uchida, Hokuto Ototake and Shigeru Masuyama	78
<i>Selective Annotation of Sentence Parts: Identification of Relevant Sub-sentential Units</i>	
Ge Xu, Xiaoyan Yang and Chu-Ren Huang	86
<i>The Kyutech corpus and topic segmentation using a combined method</i>	
Takashi Yamamura, Kazutaka Shimada and Shintaro Kawahara	95
<i>Automatic Evaluation of Commonsense Knowledge for Refining Japanese ConceptNet</i>	
Seiya Shudo, Rafal Rzepka and Kenji Araki	105
<i>SAMER: A Semi-Automatically Created Lexical Resource for Arabic Verbal Multiword Expressions Tokens Paradigm and their Morphosyntactic Features</i>	
Mohamed Al-Badrashiny, Abdelati Hawwari, Mahmoud Ghoneim and Mona Diab	113
<i>Sentiment Analysis for Low Resource Languages: A Study on Informal Indonesian Tweets</i>	
Tuan Anh Le, David Moeljadi, Yasuhide Miura and Tomoko Ohkuma	123