

24th Conference on Computational Linguistics and Speech Processing (ROCLING 2012)

**Chung-Li, Taiwan
21-22 September 2012**

Editors:

**Richard Tzong-Han Tsai
Liang-Chih Yu
Chia-Ping Chen**

**Cheng-Zen Yang
Shu-Kai Hsieh
Min-Yuh Day**

ISBN: 978-1-62748-392-6

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© (2012) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2013)

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

Proceedings of the Twenty-Fourth Conference on Computational Linguistics and Speech Processing ROCLING XXIV (2012)

TABLE OF CONTENTS

Preface	iii
Organization	iv
Program Overview	vii
Invited Speech I: How Many Multiword Expressions do People Know?	xi
<i>Kenneth Church</i>	
Invited Speech II: Deep Learning and A New Wave of Innovations in Speech Technology	xii
<i>Li Deng</i>	
 Oral Session 1: Speech Processing I	
Improved Histogram Equalization Methods for Robust Speech Recognition	1
<i>Hsin-Ju Hsieh, Jieh-Weih Hung and Berlin Chen</i>	
A Voice Conversion Method Mapping Segmented Frames with Linear Multivariate Regression	3
<i>Hung-Yan Gu, Jia-Wei Chang and Zan-Wei Wang</i>	
Acoustic Variability in the Speech of Children with Cerebral Palsy	15
<i>Li-Mei Chen, Han-Chih Ni, Tzu-Wen Kuo and Kuei-Ling Hsu</i>	
 Oral Session 2: Sentiment Analysis and Semantics	
Domain Dependent Word Polarity Analysis for Sentiment Classification	30
<i>Ho-Cheng Yu, Ting-Hao Huang and Hsin-Hsi Chen</i>	
Attachment of English Prepositional Phrases and Suggestions of English Prepositions	32
<i>Chia-Chi Tsai and Chao-Lin Liu</i>	
Associating Collocations with WordNet Senses Using Hybrid Models	47
<i>Yi-Chun Chen, Tzu-Xi Yen and Jason S. Chang</i>	
Measuring Individual Differences in Word Recognition: The Role of Individual Lexical Behaviors	61
<i>Hsin-Ni Lin, Shu-Kai Hsieh and Shiao-Hui Chan</i>	
 Oral Session 3: Speech Processing II	
Recurrent Neural Network-based Language Modeling with Extra Information Cues for Speech Recognition	75
<i>Bang-Xuan Huang, Hank Hao, Memphis Chen and Berlin Chen</i>	

<u>A Prediction Module for Taiwanese Tone Sandhi Based on the Decision Tree Algorithm</u>	92
<i>Neng-Huang Pan, Ming-Shing Yu and Pei-Chun Tsai</i>	
<u>Development of a Taiwanese Speech and Text Corpus</u>	102
<i>Tzu-Yu Liao, Ren-Yuan Lyu, Ming-Tat Ko, Yuang-Chin Chiang and Jyh-Shing Jang</i>	
Oral Session 4: NLP Applications	
<u>The Design of Chinese Character Learning System Based on Phonetic Components</u>	112
<i>Chia-Hui Chang and Wen-Pen Wu</i>	
<u>Automatic Correction for Graphemic Chinese Misspelled Words</u>	125
<i>Tao-Hsing Chang, Shou-Yen Su and Hsueh-Chih Chen</i>	
<u>Exploiting Machine Learning Models for Chinese Legal Documents Labeling, Case Classification, and Sentencing Prediction</u>	140
<i>Wan-Chen Lin, Tsung-Ting Kuo, Tung-Jia Chang, Chueh-An Yen, Chao-Ju Chen and Shou-de Lin</i>	
Oral Session 5: Machine Translation and Information Retrieval	
<u>Detecting and Correcting Syntactic Errors in Machine Translation Using Feature-Based Lexicalized Tree Adjoining Grammars</u>	142
<i>Wei-Yun Ma and Kathleen Mckeown</i>	
<u>An Improvement in Cross-Language Document Retrieval Based on Statistical Models</u>	144
<i>Longyue Wang, Derek F. Wong and Lidia S. Chao</i>	
Poster Session:	
<u>English-to-Traditional Chinese Cross-lingual Link Discovery in Articles with Wikipedia Corpus</u>	156
<i>Liang-Pu Chen, Yu-Lun Shih, Chien-Ting Chen, Tsun Ku, Wen-Tai Hsieh, Hung-Sheng Chiu and Ren-Dar Yang</i>	
<u>Skip N-gram Modeling for Near-Synonym Choice</u>	163
<i>Shih-Ting Chen, Wei-Cheng He, Philips Kokoh Prasetyo and Liang-Chih Yu</i>	
<u>Metaphor and Metonymy in Apple Daily's Headlines</u>	176
<i>Chih-Lin Chuang</i>	
<u>Phonetics of Speech Acts: A Pilot Study</u>	185
<i>Chih-Lin Chuang</i>	
<u>Convolutional Blind Source Separation Based on Sparse Component Analysis</u>	192
<i>Hsiang-Lung Chuang, Yu-Shiun Shie, Chang-Hong Lin and Jia-Ching Wang</i>	
<u>Disambiguating Main POS tags for Turkish</u>	202
<i>Razieh Ehsani, Muzaffer Ege Alper, Gulsen Eryigit and Esref Adali</i>	
<u>Automatic Time Alignment for a Taiwanese Read Speech Corpus and its Application to Constructing Audiobooks with Text-Speech Synchronization</u> ..	214
<i>Wei-jay Huang, Jhih-rou Lin, Ren-yuan Lyu, Yuang-chin Chiang, Jyh-Shing Roger Jang and Ming-Tat Ko</i>	

<u>Study on Keyword Spotting using Prosodic Attribute Detection for Conversational Speech</u>	231
<i>Yu-Jui Huang, Yin-Wei Chung and Jui-Feng Yeh</i>	
<u>Translating Collocation using Monolingual and Parallel Corpus</u>	246
<i>Ming-Zhuan Jiang, Tzu-Xi Yen, Chung-Chi Huang, Mei-Hua Chen and Jason S. Chang</i>	
<u>A Possibilistic Approach for Automatic Word Sense Disambiguation</u>	261
<i>Oussama Ben Khiroun, Bilel Elayeb, Ibrahim Bounhas, Fabrice Evrard and Narijès Bellamine Ben Saoud</i>	
<u>Applying Association Rules in Solving the Polysemy Problem in a Chinese to Taiwanese TTS System</u>	276
<i>Yih-Jeng Lin, Ming-Shing Yu and Wei-Lun Li</i>	
<u>Context-Aware In-Page Search</u>	292
<i>Yu-Hao Lin, Yu-Lan Liu, Tzu-Xi Yen and Jason S. Chang</i>	
<u>Implementation of Malayalam Morphological Analyzer Based on Hybrid Approach</u>	307
<i>Vinod P M, Jayan V and Bhadrar V K</i>	
<u>A Light Weight Stemmer in Kokborok</u>	318
<i>Braja Gopal Patra, Khumbar Debbarma, Swapan Debbarma, Dipankar Das, Amitava Das and Sivaji Bandyopadhyay</i>	
<u>Implementation and Comparison of Keyword Spotting for Taiwanese</u>	326
<i>Chung-Che Wang, Che-Hsuan Chou, Liang-Yu Chen, Yu-Jhe Li, Jyh-Shing Jang, Hsun-Cheng Hu, Shih-Peng Lin and You-Lian Huang</i>	
<u>Applications of Parallel Corpora for Chinese Segmentation</u>	341
<i>Jui-Ping Wang and Chao-Lin Liu</i>	
<u>Concatenation-based Method for the Synthesis of Engine Noise with Continuously Varying Speed</u>	356
<i>Ming-Kuan Wu and Chia-Ping Chen</i>	
<u>Collaborative Annotation and Visualization of Functional and Discourse Structures</u>	366
<i>Hengbin Yan and Jonathan Webster</i>	
<u>Improving Chinese Textual Entailment by Monolingual Machine Translation Technology</u>	375
<i>Shan-Shun Yang, Shih-Hung Wu, Liang-Pu Chen, Wen-Tai Hsieh and Seng-Cho T. Chou</i>	