2014 International Conference on Asian Language Processing

(IALP 2014)

Kuching, Malaysia
20-22 October 2014
Automatic Arabic Term Extraction from Special Domain Corpora  pg 1
AbdulMohsen Al-Thubaity, Marwa Khan, Saad Alotaibi and Badriyya Alonazi

Semantic Type Disambiguation for Japanese Verbs  pg 6
Shohei Okada and Kazuhide Yamamoto

Which Performs Better for New Word Detection, Character Based or Chinese Word Segmentation based?  pg 10
Haijun Zhang and Shumin Shi

Named Entity Recognition in Assamese using CRFs and Rules  pg 15
Padmaja Sharma, Utpal Sharma and Jugal Kalita

Building an Indonesian Named Entity Recognizer using Wikipedia and DBPedia  pg 19
Andry Luthfi, Bayu Distiawan and Ruli Manurung

“ Wanyi...Ye “  pg 23
Shuzhen Shi and Pu Li

The Usage of Zongshi  pg 26
Shuqin Shi and Kaihong Yang

Research on Building Chinese Semantic Lexicon Based on the Concept Definition of HowNet  pg 29
Hui Cao and Sen Zhang
Imperative Sentences with Assertive Mood  pg 34
Pu Li and Hao Zhao

Information Decompression of Xinjiang Travel Materials  pg 38
Kaihong Yang and Shuzhen Shi

<table>
<thead>
<tr>
<th>Session</th>
<th>SP-1: Spoken Language Processing 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date/Time</td>
<td>Monday, 20 Oct 2014 / 14:00 – 15:40</td>
</tr>
</tbody>
</table>

Acoustic Model Merging using Acoustic Models from Multilingual Speakers for Automatic Speech Recognition  pg 42
Tien-Ping Tan, Laurent Besacier and Benjamin Lecouteux

Classification of Phonemes using Modulation Spectrogram based Features for Gujarati Language  pg 46
Anshu Chittora and Hemant A. Patil

A Spectral Transition Measure based MELCEPSTRAL Features for Obstruent Detection  pg 50
Bhavik B. Vachhani, Kewal D. Malde, Maulik C. Madhavi and Hemant A. Patil

Vocal Tract Length Normalization for Vowel Recognition in Low Resource Languages  pg 54
Shubham Sharma, Maulik C. Madhavi and Hemant A. Patil

A Cepstral Mean Subtraction based Features for Singer Identification  pg 58
Purushotam G. Radadia and Hemant A. Patil
**Session CL-2: Computational Linguistics 2**

**Date/Time**
Monday, 20 Oct 2014 / 16:00 – 17:40

- **Automatic Processing of Chinese special structure ‘Ba–construction’**  pg 62
  *Yiyi Zhao*

- **Designing an Indonesian Part of speech Tagset and Manually Tagged Indonesian Corpus**  pg 66
  *Arawinda Dinakaramani, Fam Rashel, Andry Luthfi and Ruli Manurung*

- **Building an Indonesian Rule-Based Part-of-Speech Tagger**  pg 70
  *Fam Rashel, Andry Luthfi, Arawinda Dinakaramani and Ruli Manurung*

- **The Analysis on Mistaken Segmentation of Tibetan Words based on Statistical Method**  pg 74
  *Congjun Long, Yiyong Lan and Xiaobin Zhao*

- **Research on Recognition of Semantic Chunk Boundary in Tibetan**  pg 78
  *Tianhang Wang, Shumin Shi, Heyan Huang, Congjun Long and Ruijing Li*

**Session LS-2: Linguistics and Language Studies 2**

**Date/Time**
Monday, 20 Oct 2014 / 16:00 – 17:40

- **The Formation of Modern Chinese Imperative Sentence Combining**  pg 83
  *Hao Zhao and Shuqin Shi*

- **A Comparative Study of Nominal Predicate Sentences (NPS) and SHI (be) Sentences**  pg 87
  *Lijuan Zhang*
Automatic Detection of Subject/Object Drops in Bengali  pg 91
Arjun Das, Utpal Garain and Apurbalal Senapati

Logical Operative Processes of Semantic Grammar for Machine Interpretation  pg 95
Sivakumar Ramakrishnan, Pradeep Isawasan and Vasuky Mohanan

Imagistic and Propositional Languages in Classical Chinese Poetry  pg 99
John Lee and Yin Hei Kong

Effectiveness of Multiscale Fractal Dimension-based Phonetic Segmentation in
Speech Synthesis for Low Resource Language  pg 103
Mohammadi Zaki, J. Nirmesh Shah and Hemant A. Patil

Influence of Various Asymmetrical Contextual Factors for TTS in a Low
Resource Language  pg 107
Nirmesh J. Shah, Mohammadi Zaki and Hemant A. Patil

A Comparative Study on Extractive Speech Summarization of Broadcast News
and Parliamentary Meeting Speech  pg 111
Jian Zhang and Huaqiang Yuan

Development of Language Resources for Speech Application in Gujarati and
Marathi  pg 115
Maulik C. Madhavi, Shubham Sharma and Hemant A. Patil

Nonlinear Analysis of Natural vs. HTS-Based Synthetic Speech  pg 119
Hermant A. Patil and Adarsa S.
### Tuesday, October 21, 2014

#### Session: CL-3: Computational Linguistics 3

<table>
<thead>
<tr>
<th>Date/Time</th>
<th>Tuesday, 21 Oct 2014 / 10:50 – 12:30</th>
</tr>
</thead>
</table>
| **Learning Word Embeddings from Dependency Relations** | pg 123
  Yinggong Zhao, Shujian Huang, Xinyu Dai, Jianbing Zhang and Jiajun Chen |

| **A Bottom–up Method for Analyzing the Domain of a Sentence Group** | pg 128
  Xiangfeng Wei, Quan Zhang, Yi Yuan and Zhejie Chi |

| **The Refined MI: A Significant Improvement to Mutual Information** | pg 132
  Maha Alrabiah, AbdulMalik Al-Salman and Eric Atwell |

| **Polarity Detection of Turkish Comments on Technology Companies** | pg 136
  Gözde Gül Îşgüder-Şahin, Harun Reşit Zafer and Eşref Adali |

| **Discovering Linguistic Knowledge by Converting Printed Dictionaries of Minority Languages into Machine Readable Dictionaries** | pg 140
  Bali Ranaivo-Malançon, Suhaila Saee and Jennifer Fiona Wilfred Busu |

#### Session: LS-3: Linguistics and Language Studies 3

<table>
<thead>
<tr>
<th>Date/Time</th>
<th>Thursday, 21 Oct 2014 / 10:50 – 12:30</th>
</tr>
</thead>
</table>
| **An Acoustic Analysis of English Monophthongs by Tibetan Speakers** | pg 144
  Shourong An, Hui Feng Huixia Wang and Jianwu Dang |

| **Acoustic Features of Mandarin Monophthongs by Tibetan Speakers** | pg 147
  Lu Zhao, Hui Feng Huixia Wang and Jianwu Dang |

2014 International Conference on Asian Language Processing (IALP)

xxi
A Prosodic Analysis of Insertion Repair at Transition Space in Chinese Conversation  pg 151
Wei Zhang

A Computer–Aided Chinese Pronunciation Training Program for English–Speaking Learners  pg 154
Yi Qin and Guonian Wang

Use of Glottal Inverse Filtering for Asthma and HIE Infant Cries Classification  pg 158
Anshu Chittora and Hemant A. Patil

<table>
<thead>
<tr>
<th>Session</th>
<th>MT-1: Machine Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date/Time</td>
<td>Thursday, 21 Oct 2014 / 10:50 – 12:30</td>
</tr>
</tbody>
</table>

One–Expression Classification in Bengali and its role in Bengali–English Machine Translation  pg 162
Apurbalal Senapati and Utpal Garain

Extracting Parallel Phrases from Comparable Corpora  pg 166
Jiexin Zhang, Hailong Cao and Tiejun Zhao

Hybrid Approach for Aligning Parallel Sentences for Languages without a Written Form using Standard Malay and Malay Dialects  pg 170
Yen-Min Jasmina Khaw and Tien-Ping Tan

A Maximum Entropy Based Reordering Model for Mongolian-Chinese SMT with Morphological Information  pg 175
Zhenxin Yang, Miao Li, Zede Zhu, Lei Chen, Linyu Wei and Shaoqi Wang
Local Phrase Reordering Model for Complicated Chinese NPs in Patent Chinese-English Machine Translation  pg 179
Xiaodie Liu, Yun Zhu and Yaohong Jin

Session LA-1: Natural Language Applications
Date/Time Tuesday, 21 Oct 2014 / 14:00 – 15:20

A Novel Query Expansion Method for Military News Retrieval Service  pg 183
Liang-Chu Chen, Wen-Tsan Chao and Chia-Jung Hsieh

An Approach for Automatically Structuring Vietnamese Legal Text  pg 187
Thinh D. Bui and Quoc B. Ho

The Retrieval Research of Non-Adjacent Keywords in Chinese Corpus — A Case Study of “Yi...Jiu...” Construction  pg 191
Xiaoping Tan and Lijiao Yang

A Rule-Based Method for Chinese Punctuations Processing in Sentences Segmentation  pg 195
Jing Wang, Yun Zhu and Yaohong Jin

Session LR-1: Natural Language Resources
Date/Time Tuesday, 21 Oct 2014 / 14:00 – 15:20

An Extracted Database Content from WordNet for Natural Language Processing and Word Games  pg 199
Josephine E. Petralba
Automatic Acquisition of Morphological Resources for Melanau Language  pg 203
Suhaila Saee, Lay-Ki Soon, Tek-Yong Lim, Bali Ranaivo-Malan, Jovianna Juk and Enya Kong Tang

NormAPI: An API for Normalizing Filipino Shortcut Texts  pg 207
Nicco Nocon, Gems Cuevas, Darwin Magat, Peter Suministrado and Charibeth Cheng

Improving Malt Dependency Parser using a Simple Grammar–Driven Unlexicalised Dependency Parser  pg 211
Anil Krishna Eragani and Varun Kuchibhotla

<table>
<thead>
<tr>
<th>Session</th>
<th>IP-1: Language Information Processing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date/Time</td>
<td>Tuesday, 21 Oct 2014 / 14:00 – 15:20</td>
</tr>
</tbody>
</table>

Semantic Conceptual Primitives Computing in Text Classification  pg 215
Quan Zhang, Yi Yuan, Xiangfeng Wei, Zhejie Chi, Peimin Cong and Yihua Du

Category–associated Collocative Concept Primitives Extraction  pg 219
Zhejie Chi and Quan Zhang

Sentiment Classification using Enhanced Contextual Valence Shifters  pg 224
Vo Ngoc Phu and Phan Thi Tuoi

Concepts Identification of an NL query in NLIDB Systems  pg 230
Saikrishna Srirampur, Ravi Chandibhamar, Ashish Palakurthi and Radhika Mamidi