# IALP 2009 Table of Contents

**Message from General Chair** .................................................................................................................. x
**Message from Program Chair** .................................................................................................................. xi
**Conference Committees** ............................................................................................................................ xii
**Program Committee** .................................................................................................................................... xiii
**Invited Talks** .................................................................................................................................................. xv

## IALP 2009

An Articulation Training System with Intelligent Interface and Multimode Feedbacks to Articulation Disorders ................................................................. 3
*Yeou-Jiunn Chen, Jiunn-Liang Wu, Hui-Mei Yang, Chung-Hsien Wu, Chih-Chang Chen, and Shan-Shan Ju*

Research on Feature Extraction from Chinese Text for Opinion Mining ............................................................... 7
*Shanzong Zhu, Yuanchao Liu, Ming Liu, and Peiliang Tian*

Dialog-Act Recognition Using Discourse and Sentence Structure Information ........................................... 11
*Keyan Zhou and Chengqing Zong*

Multilingual Multimodal Integration of Sketch and Speech: A Generic Speech Representation Model for Spatial Description .......................................................... 17
*Lee-Na Teh and Alvin W. Yeo*

An Experimental Study on Vietnamese POS Tagging ...................................................................................... 23
*Oanh Thi Tran, Cuong Anh Le, Thuy Quang Ha, and Quynh Hoang Le*

Similarity Computing Between Simple Questions Based on HowNet and Syntactic Analysis ................................................. 28
*Zhou Zhibin, Shi Shuicai, Han Yanhua, and Lv Xueqiang*

Information Focus Synthesis Based on Question Answer Chain ............................................................... 34
*Jing Wan and Han Ren*

Stability vs. Effectiveness: Improved Sentence-Level Combination of Machine Translation Based on Weighted MBR ............................................................................................. 39
*Bo Wang, Tiejun Zhao, Muyun Yang, Hongfei Jiang, and Sheng Li*
Repetition in Mandarin Interaction: A Case Study on TV Shopping Channels in Taiwan .....................................................................................................................................................................43
   Fuhui Hsieh and Ying Liang

Building Thai WordNet with a Bi-directional Translation Method .................................................48
   Dhanon Leenoi, Thepchai Supnithi, and Wirote Aroonmanakun

Segmenting Long Sentence Pairs for Statistical Machine Translation ............................................53
   Biping Meng, Junsheng Zhou, Shujian Huang, and Jiajun Chen

Combining ILP and MLN for Coreference Resolution ....................................................................59
   Yabing Zhang, Junsheng Zhou, Shujian Huang, and Jiajun Chen

Improved Reordering Rules for Hierarchical Phrase-Based Translation ......................................65
   Shu Cai, Yajuan Li, and Qun Liu

A Hybrid Method of Chinese Prosodic Word Tagging Based on Keyword Anchor and Hidden Markov Model ....................................................................................................................................71
   Zhou Quan, Deng Pan, Liu Hongjian, Guo Defeng, and Nagamatsu Kenji

An Acoustic Comparison of Vowel Length Contrasts in Standard Arabic, Japanese and Thai ......................................................................................................................................................76
   Kimiko Tsukada

Semantic Role Based Tamil Sentence Generator ..........................................................................80
   S. Lakshmana Pandian and T.V. Geetha

Vietnamese Final Stop Consonants /p, t, k/ Described in Terms of Formant Transition Slopes .....................................................................................................................................................86
   Viet Son Nguyen, Eric Castelli, and René Carré

Foxinfo1.0: A Chinese Topic-Oriented Search Engine ................................................................91
   Ke Sun, Lei Lin, Bingquan Liu, Chengjie Sun, and Xiaolong Wang

A Bi-directional Translation Approach for Building Thai Wordnet ...............................................97
   Prissana Akaraputthiporn, Krit Kosawat, and Wirote Aroonmanakun

A Study on Semantic Role Labeling of Korean Sentence ..........................................................102
   Yude Bi and Jie Chen

Challenges in Developing Persian Corpora from Online Resources ........................................108
   Masood Ghayoomi and Saeedeh Momtazi

A Wavelet-Based Denoising System Using Time-Frequency Adaptation for Speech Enhancement ........................................................................................................................................114
   Kun-Ching Wang, Chuin-Li Chin, and Yi-Hsing Tsai

Extracting Thai Compounds Using Collocations and POS Bigram Probabilities without a POS Tagger ...........................................................................................................................................118
   Wirote Aroonmanakun

Analysis and Selection of Prosodic Features for Language Identification ..................................123
   Raymond W.M. Ng, Tan Lee, Cheung-Chi Leung, Bin Ma, and Haizhou Li

A New Way for Chinese Place Name Recognition ........................................................................129
   Quan Zhang, Yi Yuan, Nuo Li, Xiangfeng Wei, and Jianming Miao
Improving the Performance of the Link Parser ..............................................................220
   Yalamanchi Viswanatha Naidu, Anil Kumar Singh, Dipti Misra Sharma, and Akshar Bharati
Chinese Chunking Based on Coarse-Grained Part-of-Speech Features ......................................226
   Guang-Lu Sun, Yibo Xue, Zhiming Xu, and Fei Lang
Identification of Closely Related Indigenous Languages: An Orthographic Approach ..........................................................230
   Ee-Lee Ng, Alvin W. Yeo, and Bali Ranaivo-Malancon
Research on Chinese Text Summarization Algorithm Based on Statistics and Rules ...........................................................236
   Faguo Zhou, Fan Zhang, and Bingru Yang
Two-Pass Deterministic Dependency Parsing for Long Chinese Sentences ..........................240
   Ping Jian and Chengqing Zong
Refining Unit Boundaries for Mandarin Text-to-Speech Database ..................................................245
   Minghui Dong, Ling Cen, Paul Chan, and Haizhou Li
Hedge-Based Filtering Method for Semi-structure Documents ......................................................249
   Tong Wang, Chunhui Zhao, and Lei Yang
Semantic Genes and the Semantic Composition of Adjectives in Modern Chinese .......................254
   Dan Hu and Jinglian Gao
Approaches of Dimensionality Reduction for Telugu Document Classification .............................259
   P. Vijayapal Reddy, B. Sasidhar, B. Harinatha Reddy, B. Vishnu Vardhan, L. Pratap Reddy, and A. Govardhan
Automatic Chinese Essay Scoring Using Connections between Concepts in Paragraphs ..........................265
   Tao-Hsing Chang and Chia-Hoang Lee
Study on the Classification of Mixed Text Based on Conceptual Vector Space Model and Bayes ..........................................................269
   Yaxiong Li and Dan Hu
Semi-supervised Learning of Domain-Specific Language Models from General Domain Data ................273
   Shuanhu Bai, Min Zhang, and Haizhou Li
Advances in Acoustic Modeling for Vietnamese LVCSR ..........................................................280
   Tuan Nguyen and Quan Vu
A Lattice-Based Phonotactic Language Recognition System with CMLLR Adaptation and Its Implementation Issues ......................................................285
   Cheung-Chi Leung, Rong Tong, Bin Ma, and Haizhou Li
Chinese Named Entity Recognition Using a Morpheme-Based Chunking Tagger ...........................................................289
   Guohong Fu
Learning Method for Extraction of Partial Correspondence from Parallel Corpus .............................................................. 293
Ryo Terashima, Hiroshi Echizen-ya, and Kenji Araki

A Kappa Measurement of Query Consistency and Its Application ................................................................. 299
Chao Song, Muyun Yang, Haoliang Qi, and Sheng Li

Sogou Query Log Analysis: A Case Study for Collaborative Recommendation or Personalized IR ............................................................. 304
Zhitao Zhang, Muyun Yang, Sheng Li, Haoliang Qi, and Chao Song

Ranking vs. Classification: A Case Study in Mining Organization Name Translation from Snippets ................................................................. 308
Muyun Yang, Zhenyong Shi, Sheng Li, Tiejun Zhao, and Haoliang Qi

The Improved Logistic Regression Models for Spam Filtering .................................................................................. 314
Yong Han, Muyun Yang, Haoliang Qi, Xiaoning He, and Sheng Li

Automatic Acquisition of Large-Scale Academic Bilingual Parallel Corpus from the Web ................................................................. 318
Han Yong, Li Yu, He Xiaoning, Yang Muyun, and Lei Guohua

BEST Corpus Development and Analysis ............................................................................................................... 322
Monthika Boriboon, Kanyanut Kriengket, Patcharakarn Chootarakool,
Sitthaa Phaholphinyo, Somonmas Purodakananda, Tipraporn Thanakulwarapas, and Krit Kosawat

Three Algorithms for Word-to-Phrase Machine Translation ..................................................................................... 328
Le Manh Hai and Phan Thi Tuoi

A Study on Eye Movement of Korean Students Reading Chinese Texts with or without Marks for Word Boundaries ............................................................................................................... 332
Yu Peng and Liu Su

Author Index ....................................................................................................................................................... 337